

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2002, 12:34:06 ; Search time 5062.66 seconds

(without alignments)
11825.456 Million cell updates/sec

Title: US-09-357-273a-1

Perfect score: 3629

Sequence: 1 ccgcgtcgcgcgcgcgtca.....caggagccagcccgatcc 3629

Scoring table: IDENTITY_NTC

Gapop 10.0 , Gapext 1.0

Number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vi:*

30: em_htgo_hum:*

31: em_htgo_inv:*

32: em_htgo_rod:*

33: em_htg_hum:*

34: em_htg_inv:*

35: em_htg_rod:*

36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3629	100.0	3629	9 AF059198	AF059198 Homo sapi
2	2503.2	69.0	3976	10 AB031332	AB031332 Mus muscu
3	801.4	22.1	154460	2 AC025362	AC025362 Homo sapi
4	798.2	22.0	162473	9 AC005803	AC005803 Homo sapi
5	675.2	18.6	2874	9 AB047079	AB047079 Homo sapi
6	655.6	18.1	2769	6 AX056432	AX056432 Sequence
7	624	17.2	2887	10 AF071777	AF071777 Mus muscu
8	548.2	15.1	70508	2 AC016784	AC016784 Homo sapi
9	281.6	7.8	239140	2 AC013350	AC013350 Mus muscu
10	276.8	7.6	577	11 G10027	G10027 human SRS C
11	266	7.3	87944	2 AC023662	AC023662 Homo sapi
12	246	6.8	132830	2 AC011051	AC011051 Homo sapi
13	228.6	6.3	155751	3 AC007711	AC007711 Drosophi
14	228.6	6.3	158511	2 AC014325	AC014325 Drosophi
15	228.6	6.3	175917	3 AC008142	AC008142 Drosophi
16	203.2	5.6	39750	3 CEC41C4	CEC41C4 Caenorhadi
17	193.2	5.3	60004	2 AC013601	AC013601 Homo sapi
18	193	5.3	60004	2 AC013601	AC013601 Homo sapi
19	168	4.6	132830	2 AC011051	AC011051 Homo sapi
20	156	4.3	87944	2 AC023662	AC023662 Homo sapi
21	128	3.5	70508	2 AC016784	AC016784 Homo sapi
22	109.6	3.0	211967	2 AC013368	AC013368 Homo sapi
23	103.2	2.8	507	1 AF315626	AF315626 Haloferax
24	101	2.8	135751	2 AC012317	AC012317 Homo sapi
25	101	2.8	150030	2 AC008870	AC008870 Homo sapi
26	101	2.8	160212	2 AC009043	AC009043 Homo sapi
27	101	2.8	183988	2 AC012185	AC012185 Homo sapi
28	89.6	2.5	405	6 AX072804	AX072804 Sequence
29	77	2.1	19000	8 SPAC167	SPAC167 S. pombe c
30	76.2	2.1	5420	8 YSCDCPK	YSCDCPK Saccharomyc
31	76.2	2.1	39203	8 YSCCH9205	YSCCH9205 Saccharomyc
32	74.6	2.1	3961	8 SCIRE1DNA	SCIRE1DNA Saccharomyc
33	74.6	2.0	197232	2 AC073946	AC073946 Mus muscu
34	74.2	2.0	3121	8 AB031396	AB031396 Oryza sat
35	68.6	1.9	2858	9 MFAPA04A	MFAPA04A M. fascicula
36	67.6	1.9	1365	9 BABATV	BABATV Baboon apol
37	65.8	1.8	660	6 CNS01B3G	CNS01B3G Botrytis
38	65.8	1.8	1968	10 MMU04663	MMU04663 Mus muscu
39	65	1.8	162869	9 AC090645	AC090645 Homo sapi
40	64.4	1.8	5120	6 AX067148	AX067148 Sequence
41	64.4	1.8	5120	9 AF152102	AF152102 Homo sapi
42	64.2	1.8	166874	33 AC021022	AC021022 Homo sapi
43	64.2	1.8	168778	9 AC090886	AC090886 Homo sapi
44	64.2	1.8	168590	9 AC090004	AC090004 Homo sapi
45	63.8	1.8	2126	8 AB042550	AB042550 Oryza sat

ALIGNMENTS

RESULT 1

AF059198 3629 bp mRNA PRI 11-JUL-1998

LOCUS Homo sapiens protein kinase/endoribonuclease (IRE1) mRNA, complete cds

ACCESSION AF059198

VERSION AF059198.1 GI:3300093

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3629)

AUTHORS Tirasophon,W., Welihinda,A.A. and Kaufman,R.J.

TITLE A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Irelp) in mammalian cells

JOURNAL Genes Dev. 12 (12), 1812-1824 (1998)

MEDLINE 98301437

REFERENCE 2 (bases 1 to 3629)

LOCUS AB031332 3976 bp mRNA ROD 12-JAN-2001
DEFINITION Mus musculus Irel alpha mRNA for protein kinase/endoribonuclease(Irel) alpha, complete cds.
ACCESSION AB031332
VERSION AB031332.1 GI:12083699
KEYWORDS protein kinase/endoribonuclease(Irel) alpha.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Miyoshi, K., Katayama, T., Imazumi, K., Taniguchi, M., Mori, Y., Hito, J., Yui, D., Manabe, T., Gomi, F., Yoneda, T., and Tohyama, M.
TITLE Characterization of mouse Irelalpha: cloning, mRNA localization in the brain and functional analysis in a neural cell line
JOURNAL Brain Res. Mol. Brain Res. 85 (1-2), 68-76 (2000).
REFERENCE 2 (bases 1 to 3976)
AUTHORS Miyoshi, K.
TITLE Direct Submision
REFERENCE Submitted (21-AUG-1999) Ko Miyoshi, Osaka University Medical School, Department of Anatomy and Neuroscience, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
COMMENT (E-mail:miyoshi@anat2.med.osaka-u.ac.jp, Tel: +81-6-6879-3221, Fax: +81-6-6879-3229)
FEATURES Sequence updated (01-Sep-1999).
source Location/Qualifiers
 1..3976
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 121..3054
 /gene="Irel alpha"
 121..3054
 /gene="Irel alpha"
 /codon_start=1
 /product="protein kinase/endoribonuclease(Irel) alpha"
 /protein_id="BAB20901.1"
 /db_xref="GI:12083700"
 /translation="MPARWLILLALLLPSPGSGFRTSTLTPELLFVSTLDGL
 HAVSRGTGSIKWLKEDPVLYQVTHNEEPAPLPDPEKSLTGLGNNNGTLKLPFTI
 PELVOLSPCRSDGLIYMGKODIYVLDLGEKOOTLSSAPSLISLTLGRT
 TEYITMYDKTRKRLRMNATYFDYASLPEDVDXKMSFPGSCGIVYTVSESDY
 LMIQNTASPVYAVYWGQEVLRKVVHINAVETLRLTFPMSGEVGRITKWKIPFERET
 EAKSLVLTPLVYKISTSLYASPSMVEHVAVPKSTPLPLEGQTDVITGDGEC
 VTPSTDLKFDGLGSKLNLRYNWLIGHHEPLSALMLERFPNLPKHEHY
 IPADSEKSEFEVINIVGOTSDNPTPTVSQDYEEKLARAPEAPVDMLKDMATII
 LSTPLGVNVAIITTYPLSVHOOROLOHPOKLEKILLOOOLPFPBGDLNDP
 BELSSGPSESSSSSPSPRASNSHSPSSASRACTSPSLGDDDDDETRMYIV
 CKISCPKDVLCGAGETIYVKGFEDNRVAKRILPECFSTADREVQLRSEDEHPN
 VIRYCTEKDRFOYITALELCATLQEVYEQDFNLGLSEPTLLRHQTTSGLAHLST
 NIVHRDLRPHNILLSPNPNHGRIKAMISDFGKILAVGRHSRSRGVPTGEMHAP
 EMLSBDCDNPTIYVDFISAGCFYVISEGNHPGSKLQORANILLGNCDFHAP
 KHEDVIARLEIKIAMDPOORSAKHVLEKHPFMSLEKOLQDFDVSRIKEALDG
 PIVRLERGGRAVAVKMDRENTVPLQDTRKFRYTKGSGSVRLRAMENKHHREL
 PIVEVQETLGSIPDDEVRYRTSNFPHLLSHHTTQAMELCRHERLFTYTHHEPTPEPPV
 IFPAL.
BASE COUNT 937 a 1098 c 1074 g 867 t
ORIGIN
 Query Match 69.0%; Score 2503.2; DB 10; Length 3976;
 Best Local Similarity 84.7%; Pred. No. 0;
 Matches 2935; Conservative 0; Mismatches 498; Indels 33; Gaps 10;

Db 210 GCCTGAACCTCTGTTGTTGTCTGACCTCGATGGAAGCTTGATGCTGTAGCAAG 269
 Qy 240 gaaagctcaataaataatgagctttaaagaagatccagcttcaggtcccaacatgt 239
 Db 270 GACGGGCTCATCAAGTGAAGTCTTAAAGAAAGATCCAGTCTCAGGTCCCAACACGT 329
 Qy 300 gaaagagctgccttcctccacatccataatgtagagctgtatagcttgaaagca 359
 Db 330 GGAAGACCGGCTTCTCTCCACATGCCATATGATGGCAGCTGTACACATTGGAGGCA 389
 Qy 360 gaataatgaagagctcagaaatcccttaccatcccaatctggtcaggaatcc 419
 Db 390 GAACAACGAAGGCTGTGACGAATCTCCCTTACCATCCAGAAATGGTTGGTCCGCTCC 449
 Qy 420 atgcgaagtcagaaatgaaatccctacataggttaaaagcaagacatctgagt 479
 Db 450 ATGCCGAAGTTCAGATGGAATCCTTACATGAGTAAAGCAAGATATTTGGATCTTAT 509
 Qy 480 tgaactctcagcaggaagaagcagcagacttgcatcagccttgacagatagctctg 539
 Db 510 CGACCTCTGACGTGGGAGAGACGAGACTTTGTGATGCGCTTGTGATGTCCTG 569
 Qy 540 cccataaactctctctgtatcttgaggogaagaatacacatcaatgtagaac 599
 Db 570 CCCATCAACTCTCCCTTATATCTTGACGAGACGAATACCATCATCATTATGACAC 629
 Qy 600 caaaccagagctcagcaggaatcccaactcttgatcagcagcctcagctgctg 659
 Db 630 CAAGACCCGGGAGCTCCGCTGGAATGCCACTTTTGTGATATGACGCTCATGCGCGA 689
 Qy 660 gaaagagagagactacaagaatgccactctgtgtccaaatgtagatgagctgtgtgac 719
 Db 690 AGACGAGCTGACTACAAAGATGCCACTTGTGTCATATGGCATGAGACTGGTGTAC 749
 Qy 720 tgtgagagtgaaatctgaggagctctgtgtgataccaaatacgcctccctgtgtgac 779
 Db 750 TGTGAGACAGTAATCTGGGATGTCTGTGATCAAAATATGCTCTCTGTGTGTC 809
 Qy 780 ctcttatgtcagcagcagagaggtctcgaagaagtgatgacataatgtctgtg 839
 Db 810 CTCTACGTCTGGCAGGGAGAGTCTCTGAAGAAGTGTGACATACACTTGTCTGTGA 869
 Qy 840 gaccctgcgtatctgaccttacccttgcgtggaagtgagggcgatcaacaagtgaag 899
 Db 870 GACTTACGCTACTTACCTTATGCTGTGAGGAAGTGGGGCATCACCAAGTGAAGTA 929
 Qy 900 cccgtctcccaagagaacagagccaaagcaagctgagccactctgtatgtggaa 959
 Db 930 TCCATTCCCAAGAGACAGAGGCCAAGACCTTACGCTTATGTTGGGA 989
 Qy 960 atactctacagcctctatgacctccctcccaatggtgatacagaaggggtgtgtg 1019
 Db 990 GTATTCCACACGCTCTATGCTCTCTCTCAATAGTGTGCATGAGGGGCTTGTCTG 1049
 Qy 1020 ccgcygacacacatctccttctgtaagagcccccaagactgagtcgtacacatcg 1079
 Db 1050 TCGAGGACAGACTCTCTCTTGTGCTGGAAGGCCCCAGACAGATGGCGTATGGA 1109
 Qy 1080 caagggagagtggtgatacagcccaagcagcagcgtcaagttgatacccgagta 1139
 Db 1110 CAAAGGAGAGTGTGATCATCTCCAGACAGCTCAAGATTGAGCTTGAGACTCAA 1169
 Qy 1140 caagaacagctcaactacttgaggaattactgctctgtagagacacataaagccc 1199
 Db 1170 GAAGACCAAGCTAATCTACTTGAAGAAATTAAGCTTCTCATAGGACACCATATGA 1229
 Qy 1200 actgtctgcttaccacaagatgctgagagatctcccaacaatctacccaacatcg 1259
 Db 1230 TCTGTCTGATCATCAACAGATGCTGAGAGATTCTTAACAACCTGCCAACAATGAGA 1289
 Qy 1260 aaatgtatctctgtcgtatcagaagaaagagcttggaaggtatatacactgtgtga 1319

Db 1290 AATGTGATTCCTGCTGATTCAGAAAAAGAGCTTTGAGNAGTTATCAACATGTTGG 1349
Oy 1320 ccaagcttcagaagaacgacctaccggtgtctcgagatgtggaggaacccggcca 1379
Db 1330 CGAGCTTCAGACACACACGACACCGATCTCAGAGTGGAGGAGAACTCGTCG 1409
Oy 1380 tgcctctgcggccggagcccccgtgactccatgtcttaagagacagctaccat 1439
Db 1410 CGCCCTCCCAAGCTGAGGCCCCGGTGACTCCATGCTCAAGAGCATGGCTACATTAT 1469
Oy 1440 ccttgagaccttctgtgtatgtgtgtggtgtggtctcatcatcaccatccccagcat 1499
Db 1470 CCGAGCAGCTTCCGCTGTGGTGGATGGGTGGCTTCATCATCTTACCCCTGAGCGT 1529
Oy 1500 gcatcagaagacgagctccagccagcaagcttcagaagaacaggaggaagaatcagct 1559
Db 1530 GCATGAGAGGCTCAGCTCCAGCACCACAGTCCAGNAGGAGCTGGAGAAATTCAGCT 1589
Oy 1560 cctgcaagcagcagcagcagctgtcccttccaccaccttgagagacagctcagacg 1619
Db 1590 CCT-----GCAGCAGCAGCAGCTGCCCTTCCACCACAGGAGACCTTACCAGAGCC 1643
Oy 1620 cgaagctcttgacaacgtctgcccgtactcagaagagctcgggacacagccccagac 1679
Db 1644 TGAGTTCCTGGATTCATCTGGCCCTCTCCAGAGAGCTCTGGCACACAGCCCGCCAGCC 1703
Oy 1680 gtcccccagggcctccaccacactgtctgtccggcagagctctgtccaaagctggcag 1739
Db 1704 ATCCCCAGAGCTCCAAACCACTCCCTCCACCACCCAGCAGCTGCTCCAGGCCCCGAC 1763
Oy 1740 cagccccctcttgaaacaagaacgatlgaagaaacacagcgtgtgtatgttgaa 1799
Db 1764 CAGCCCTCTCTGGAGCAGATGATGAGATGAGAAACAGATGATGTTGGGA 1823
Oy 1800 aattcctctgtcccaagatgtctgtggcagatgagagctgagggacaaatgtgtacg 1859
Db 1824 AATTTTCATCTGCCCAAGAGATGCTGTGGTGGAGCTGAGGAGCAAAATGTATACA 1883
Oy 1860 gggcattgttgaacacgagcgtgtgagagagatcccccagaggttttagct 1919
Db 1884 AGGTATGTTGACACACAGATGTGGCCGTGAAGAGATCCCTCGATGTGTTAGCTT 1943
Oy 1920 cgcagacggtgagctcagctgtgtgagaaatcgatagcaacccgaaacgtgtacgcta 1979
Db 1944 TGCCAGCGGTGAGGTCAGCTGCTGAGAAATCAGACGAGCACCAGAAATGTATCGCTA 2003
Oy 1980 ctctgcacggaagagacggcgaatctccagtaactgtccatcgagctgtgtgcagca 2039
Db 2004 CTTTTCACAGAGAGAGACCGGACGTTCCAGTACATTTGCTATCGAGCTGTGACGCCAC 2063
Oy 2040 cctgcaagatlatgtgagcagaagagcttgcgcatctcgccctggagcccatcact 2099
Db 2064 CCTACAAAGATATGTGGAGCAGAAAGACTTTGCCCACTTGCCCTCGAGCCCATACCT 2123
Oy 2100 gctgcagcagcaccctcggcgtgtgcccacctcactccctcaacatcgttaccagaa 2159
Db 2124 GCTTCATCAGACACACTGAGCGCTGGACACCTGCATTTCTCAACATTTGTACAGANA 2183
Oy 2160 cctaaagcacaacatcctcatctcatgtgccaaatgtcacacgycagaatcaagccat 2219
Db 2184 CCGTGAAGCCCAACATTTCTCTCTCCATGCCCCAACGCACTTGGCAGAGATTAAGCGAT 2243
Oy 2220 gatcctcagacttgcctctgcaagaagcgtgcaatgtggcagacacagttcagccg 2279
Db 2244 GATCTCTACTTTGGCCCTCTGCAAAAGCTGGCAATGGGCAAGCAAGTTTACGCCGCCG 2303
Oy 2280 atctggggtgtcctgacagaagcgtgatacgtccacagatgtgtgagcyaagactgtaa 2339
Db 2304 TTCAAGGGGTACTCGCATGGAAGGCTGATCGCCCAAGATGTGAGAGAGACTGTAA 2363
Oy 2340 ggaagacacctactcaagcgtgtgagatcttctgcagagctgtgcctttactaagtggt 2399
Db 2364 GGACAAACCTTACTACAGCGGTGACATCTTTTTCAGAGCTGTGTCTTTACTATGTCTAT 2423
Oy 2400 ctctggagcagcaccctcttggcaagtccttgacggcagcggccaaacatctcctgg 2459
Db 2424 CTCTGAGGGCAACCATCTTTTGGCAATCTTTGACGCGGACGACCATCTCTCGG 2483
Oy 2460 tgcctgcagccttgcctgtcaccagagaagacagaagcgtcatctgcagaaat 2519
Db 2484 CGCTGCAACCTTGAAGCTGTTCCACTCAGACAAAGCATGAGAGCGTATGCTCGTAAT 2543
Oy 2520 gatagaagaatgatgtgagatgtatcctcagaagccccctcagcgaacgctgtctaa 2579
Db 2544 GATAGAGAAATGATTTGATGATGCCAGCAGCGTCCCTCTGCAAGCAGCTCTGA 2603
Oy 2580 aacccgtctctgtgagctagagaagagagcttcagttctccagagctgtgagcag 2639
Db 2604 ACACCCCTTCTTGTGAGCTGGAAGCTGGAAGCAGCTCCAGTTTTCAGAGATGAAGTACG 2663
Oy 2640 aatagaagaagaaatccctgtgatgtgcctgtgaagcagatlagagaagcggagagc 2699
Db 2664 AATAGAAAAGAGGCTTGGAGCGTCCATCTGACGGAGTTGGAGAGGCGGAGAGC 2723
Oy 2700 cgtgtgaaagatgtgactgtgaggaagaacatacctaccctccagaaagactgtctaa 2759
Db 2724 TGTGTCAGATGAGCTGCGGGAGAACATCATCTCCCTCGCAGACAGATCTCGCA 2783
Oy 2760 attcagaccctaaaggtgtgtcagagatccctccagccagcaatagaaga 2819
Db 2784 ATTGAAACCTTCAAAAGGCGCTCTGTGAGAGACTCTCCGACGCTAGAGAAACAGAA 2843
Oy 2820 gaaccactacccggagcgtgtcagaggtgtcgagagagcgtgaggacccctccagca 2879
Db 2844 ACACCACTACCGGAGCTCCCGGTGAGGTTGAGAGAGCGTGGGCTCCATCCCGATGA 2903
Oy 2880 ctctgtgtactctcaagtlcctgccttcccaactccctgcagacacactaccgggcat 2939
Db 2904 CTTTGTGGCTACTTACCTCCGCTCCCACTCTCTCTCAACCTACCAAGCAT 2963
Oy 2940 ggaagctgtcagacagagagactctccagccactcactcagaccccccagagc 2999
Db 2964 GGAGCTGTCAACATGAGAGACTCTTTCAGACACTACTGTGGACAGCCACAGAAC 3023
Oy 3000 ccaagcccaagtactcagacgcccctctgagagagcggccctctgtctgtgtgccc 3059
Db 3024 CCAGCCTCAGATGATTCATATGATGCTGAGCTGAGGAG--CCCTGTGTTGTGGTGGC 3082
Oy 3060 ccaagctgtgacgagagcgtgtgtccacaaatgaagctgtgtgtcccgcttggag 3119
Db 3083 CCAATTAATGACCATGGGCGCCGATCTCTGATGATGTTGTTGCTCTGGGATTAGCAG 3142
Oy 3120 ggaagcagagcttccaaacaaagctgtgagctgtgctgtcagcccaagagac 3179
Db 3143 GAAGACTAAGCTTCCCAATCAAGTGCTGTGAGCTG--CTGATCTCGACGCAACAGAGAT 3201
Oy 3180 agtctgaccccaagaaatgtggaagatgtgcccctgtgtgactacaaggaactgtgaa 3239
Db 3202 AACGCTGATCTTACAGACGAGGAGAAAGTGTCTTCATGACTTACAGAGCTGAGGAGA 3261
Oy 3240 tgcgtgcccacaagacgtatcagatgagtgatgtctcaagagag---gctcagaagaag 3296
Db 3262 TG--TGCCCTGAACCTTGTATGTGAAGAGCTCTGAGAAAGGACGCTGCCAGAGCGTG 3320
Oy 3297 cgcagtagaccaccacacactactagataaactgtctc-----agaacttlt 3347
Db 3321 CAAGAGAAACACATCAGAGCTTTCACGCCGATGAGCTTGCTCCACTTCTTTCTTCTA 3380
Oy 3348 aaatcctgtctaatgtcagtlctacagccttcaagaaaggaaggaagaaatcgttaca 3407
Db 3381 AAATTCCTGTGGATGGAATTTTGGGGGCGCT---TCAGTGAAGTGTGAGAGAACTGTGT 3437
Oy 3408 ttgtcgtgctgtgagacagctagcgtgagatgtacaaatgtacagcttcaagtgaga 3467
Db 3438 TTTGCTGCATGTGTGAAGACGCTGTGGGTATTTGATGTGTGACCTCTGAT--AGA 3495

LOCUS AC005803 162473 bp DNA PRI 30-OCT-1998
DEFINITION Homo sapiens chromosome 17, clone hRPK_214_C_8, complete sequence.
ACCESSION AC005803
VERSION AC005803.1 GI:3810668
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hRPK_214_C_8
JOURNAL Unpublished
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C.,
Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K.,
Donnellan,L., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R.,
Gage,D., Gardyna,S., Geraigery,K., Grant,G., Hages,B., Hearford,A.,
Herena,L., Horton,L., Howland,J.C., Jacotot,L., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Macdonald,P., Margulis,N., McEwan,P.,
McGurt,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J.,
Mychalecyy,J., Nahr,R., Naylor,J., Niof,M., O'Connor,T.,
O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,
Severy,P., Stange-Thomann,N., Stillwell,J., Stojanovic,N., Stone,C.,
Subramanian,A., Testfave,S., Tichovolsky,N., Torunella-Miller,I.,
Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D.,
Ye,W.J., Zhao,J. and Zody,M.
JOURNAL Direct Submission
AUTHORS Submitted (30-OCT-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 162473)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C.,
Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K.,
Donnellan,L., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R.,
Gage,D., Gardyna,S., Geraigery,K., Grant,G., Hages,B., Hearford,A.,
Herena,L., Horton,L., Howland,J.C., Jacotot,L., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Macdonald,P., Margulis,N., McEwan,P.,
McGurt,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J.,
Mychalecyy,J., Nahr,R., Naylor,J., Niof,M., O'Connor,T.,
O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,
Severy,P., Stange-Thomann,N., Stillwell,J., Stojanovic,N., Stone,C.,
Subramanian,A., Testfave,S., Tichovolsky,N., Torunella-Miller,I.,
Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D.,
Ye,W.J., Zhao,J. and Zody,M.
JOURNAL Direct Submission
AUTHORS Submitted (30-OCT-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 30, 1998 this sequence version replaced gi:3805774.
All repeats were identified using RepeatMasker. Smlt, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES
source
1..162473
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPCT-11 human BAC library"
/map="17"
/chromosome="17"
364..674
/rpt_family="AluY"
complement(808..904)
/rpt_family="MIR"
956..1258
/rpt_family="AluSx"
1979..2265
/rpt_family="AluSg1"
complement(2801..2901)
/rpt_family="L2"
repeat_region
complement(2857..2918)
/rpt_family="MIR"
complement(3650..3952)
/rpt_family="AluY"
repeat_region
complement(3972..4008)
/rpt_family="(TGA)n"
complement(4011..4189)
/rpt_family="L2"
4194..4241
/rpt_family="(TA)n"
complement(4242..4532)
/rpt_family="AluSg"
4533..4550
/rpt_family="(TA)n"
4838..5144
/rpt_family="AluSg"
5301..5327
/rpt_family="AT-rich"
complement(5639..5813)
/rpt_family="AluJo"
complement(5814..6104)
/rpt_family="AluSp"
complement(6105..6240)
6858..7027
/rpt_family="AluJo"
7052..7121
/rpt_family="MIR"
complement(7136..7433)
/rpt_family="AluSx"
7756..8036
/rpt_family="AluSx"
8039..8073
/rpt_family="(TAA)n"
12325..12415
/rpt_family="FAM_C"
complement(13947..14000)
/rpt_family="MER91A"
14856..14954
/rpt_family="L1PB2"
15756..15813
/rpt_family="MIR"
complement(15818..16128)
/rpt_family="AluJo"
complement(16446..16548)
/rpt_family="L1"
complement(16549..16842)
/rpt_family="AluSg"
complement(16843..17256)
/rpt_family="L1"
17257..17415
/rpt_family="L1PB2"
17416..17716
/rpt_family="AluY"
17717..18053
/rpt_family="L1PB2"
18378..18683
/rpt_family="AluJo"
19756..19797
/rpt_family="(CA)n"
20115..20214
/rpt_family="(CA)n"
complement(22960..23000)
/rpt_family="(CA)n"
23001..23041
/rpt_family="(GA)n"
23044..23180
/rpt_family="MIR"
23918..23941
/rpt_family="(GA)n"
complement(24252..24311)
/rpt_family="MER91A"
complement(24845..25154)
repeat_region

[illegible][illegible]

1652 agagctcgagaccagcagcccccagcgtcccccagggcctcccaaccactcgtctcgt 1711
1479 -----CAMTCGCCAGATGCCCCACTCCCTGCATCTGGGGGTCACCTGAGAGCAAGAA 1532
1712 ccggagctctcgctcccaagctcgcagcagccctcccttgaaacaagaatgagatg 1771
1533 GAGGCTTCAAAATCCCTCGAAGCAAGCCAGCCACTC-----GACGACCTGAAAGCTG 1585
1772 aggaacacagcgtggtgaagctgggaaatctcctctcgtcccaagagatgctcgggac 1831
1586 AGCAACTCACC-----TAGTGGGGAAGATTCTCTTCATATCCCAAGACGTCGCGGC 1639
1832 atgagctcgaaggcacaatctgtatccgggacatgttgacaacccgcagcgtgcgtga 1891
1640 ACGGGGACAGGCGGACTTTCGTTTCCGGGACAGTTGAGGAGAGAGAGCGCTGTCA 1699
1892 agagatctctcccgagctttagcttcgcagacccgtgaagctccagctgttgagaaat 1951
1700 AGCGCTCTCCGCGAGTGTGCTTGGCTTGGCGGAACTTCAACTGCTGACAGAGT 1759
1952 cggatgaagaccggaagctgatalccgctactctcgcagcggaagagacggcaatccagt 2011
1760 CTGACAGGACACCCCAATGTGCTCCGCTACTTCTGCACGAGGGGAGCCCACTTCCACT 1819
2012 acattgcaacagcgtgtgtgcagccaccctgcgaagatgtagtgagcagaagacttg 2071
1820 ACATTGCCCTGAGCTGCTCCGCGCTCTCTGACAGAGTACCTAAGAAACCCGACCTGG 1879
2072 cgcactcggcctcgtgagccactacactctcgtcgcagcagaaccccccgcggccacc 2131
1880 ATCGGGGGGCTTGAGACCCCGAGGTGTCTTCAGACGCTGATGTCTGGCTCGGCCACC 1939
2132 tccactcctcacaacatcglttcaacagaactaaagcacacacatccctacatcctcgc 2191
1940 TGCATTCTTTACATAGTACGACCGGACCTGAGCAGCAGCAATATTCATCACCGGCG 1999
2192 ccaatgcacacggcgaagatcgaagcgaatgctccgacttgcctctcgcgaagacttg 2251
2000 CTGACACCCAGGCGTGGGCGAGGTGCTTTCAGACTTCGGCTCTCCCAAGAACTCC 2059
2252 cagtgcgagacacagcttcaagccgcgaatcctggagtgctcgtcgcagagaagctgatcg 2311
2060 CTGCTGGCGCTGTAGCTTACACCTCCACTCCGCGATCCCGGACGAGGAGGCTGATGG 2119
2312 ctccagagatgctcgcagagaacatgtaagagaaacctactacacagctgtagaatcttt 2371
2120 CCGCCGAGCTCTGACAGCTCTGCCACCAACAGTCTACAGCGCTGCGACATCTCTCT 2179
2372 ctgcagagctcggtctttaaactcgtgtctcgtgagggcagcacccttttgcaagctcc 2431
2180 CTGACAGGCTGCGTGTCTACTACTGCTGCTTGTGTGGCCACCCCTTTGAGAGACAGTC 2239
2432 tgcagcgagagcacaacatcctcctggtgctcgcagccttgcagcttgcacccagaga 2491
2240 TTTATTCGCCAGGCAAAATCTCTACAGAGGCTTCCCTGTGTGGCTACCTCGAGGAAGAG 2299
2492 agcagaagacgtlcatlgtcacgagaatltgaatagagaatgtatgcgatgtatccctaga 2551
2300 TCCACGACAGAGGTGTTGCCCGGAGCTGTGACGCCATGTTGAGCCACTACTGCCACAG 2359
2552 aacgccccctcagcgaagacgctcacaacaccgttctcgtgggacctagaagagcagc 2611
2360 CACGCCCTCTCCCGCCGAGGTGGCTGGCCACCCCTCTTTTGGAGACAGCCCAACAC 2419
2612 tccagcttctcgaagacgtgagcgacagaaatagaagaagaatccctcgtgaatggccgacg 2671
2420 TCCAGTCTTCCAGATGTGCTACGTGCTGTGAGAGAGAGTGTGAGCGAGGCCCTTA 2479
2672 tgaagcagtltagagagagcgagagagccgtgtgtgaagaatgagcttgcgtgagaacatca 2731
2480 TGAAGGCAATGAGAGGAGGAGGCTGACGCTGTCCGGGACAACTGCGACGAGACATCT 2539
2732 ctgacccccctcagagacagactgtcgtaaatltagaactataaagtggttctgtltagag 2791

2540 CCATCCCGCTGAGATAGACCTGAGAAAGTTCCGCTCTATAGGGGACATCAGTGCAG 2599
2792 atctctccgagccataggaataaagaagcaacatcaccagagctgcgtcagaagctgc 2851
2600 ACTGCTCCGTGCTGTGAGAAACAAGAACACACTACAGGAGAGCTCCAGTGTGAGGTGC 2659
2852 gggagacgtctgggagaccctcccgagactcgtgtgtaactcaagctcgtctccccc 2911
2660 GACAGCACTTGGCAAGTCCCTGACGGCTTGTCCAGTACTTCAAAACCGCTTCCAC 2719
2912 acctctcgcaacacactacacggcgacatgagctgtgcaagccaagagacactccagc 2971
2720 AGCTCTCTCTCCACACGACCGAGTATGAGGAGTGTGCGCTGTAGAGACCTTCTGTC 2779
2972 cctactacttccaca 2987
2780 CCTACTACCCACACAGA 2795

RESULT 6
AX056432 2769 bp DNA PAT 13-JAN-2001
LOCUS Sequence 76 from Patent WO0073469.
DEFINITION AX056432
ACCESSION AX056432.1 GI:12229139
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2769)
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 76 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source Location/Qualifiers
1..2769
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 553 a 895 c 809 g 512 t
ORIGIN
Query Match 18.1%; Score 655.6; DB 6; Length 2769;
Best Local Similarity 55.5%; Pred. No. 7.1e-124;
Matches 1607; Conservative 0; Mismatches 1094; Indels 195; Gaps 10;

95 ccattccggcccgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgatt 154
23 CGAGGCCGTGGCCCGCGCTGGGGCTTCACACTCCAGTTCGCGGCGCTGCTCGGGACGC 82
155 ttggaagtlaccagcagtgagcgttccgtgaacctgtgtgtgtgtgtgtgtgtgtgtgtgt 214
83 TGAGTCCACAGGTTGATCTCTACAGGCCAGAGAACCTCTGCTGTGTCACCTTGATG 142
215 gaagttgcatgtgtcagcagaagagacagctcaatcaaatgagactttaaagaagatc 274
143 GAAGTCTCCACACCACTAACCAAGCAGAGGAGACCTGAAGTGAAGTCTGAGAGATGATC 202
275 cagtcctgaggtlcccaacacatltggaagagccttgccttctcccaagatccctaagt 334
203 CCGTACTGCAAGAACCAATGTACGTACAGAAATGGCTTTCTCTGTGACCCAGCAGATG 262
335 gcagcctgtatagctgtggaagcaagaataatgaagagcctgcagaaacttcccttlacca 394
263 GCAAGCTGTACATCTTGGGGAGCCCAAAACAACAGGAGATTATGAAGATGCATTACACA 322
395 tccagaatltgtcagagcatcccaatgcgaagttcagaatgaaatccctacatggtta 454
323 TCCCTAGCTGTGTCATGCTCTCCCTCGGACGCTGTGATGGGCTCTTCTACACAGGCC 382
455 aaagcagagacatcgtgtatgtatgtatgtacctcctgcagcggaagaagcagcagacttgt 514

Db	383	GGAAAGCAGGATGCTGGTTTGTGTGTGGACCCCTGAGTCAAGGAGAACCCAGATGAGACTGA	442
Qy	515	catgcgaccttgcagatagtctctgcgaatcaacatctctctctgtatctctgtggcgaaacag	574
Db	443	CCACAGAGGGT-----CCGTCACCCCCCGGCTCTAATTTGGCCGAAC	487
Qy	575	aataacacatcaacatgtatgaacaccaaaccgagagctccggtggaatgtccactact	634
Db	488	AGTATAGGGTACCATCATGACCCAAAGGCCCAACCCCTGGCGGTGGAACACACCACTTCC	547
Qy	635	ttgactatgcggctcaactcgtccgaagaagaaggactcaagaatgtccacttgyt	694
Db	548	GGCCGTACTCAGCCCCCCCCCATGGATGGCTCACCTGGAAATACATGACCACTGGCGT	607
Qy	695	ccaatgtatbgygcgtggtgtgtactgtgacagtgaatctgaggagcgtccctgtgaaacc	754
Db	608	CTCTCGGGATGGCGCTCTGCTCAGTGTGACCCAGGAAGCGGAGGGTGTGTGACAC	667
Db	755	aaactcaagctcccccgtgtgtgacctttatgtctgagagggaggtctgagaag	814
Db	668	AGGACCTGGGGGTGCTGTATAGGGGTCTACACCTGGACACAGAGCGGCTGGCCACAGC	727
Qy	815	tgatgcacatcaatgtctgcgtgtgagacccctgcgcatacttgacatctatgtcttggag	874
Db	728	TGGCGCATCTCAGCCTGGCTCGAGACACTGTGATTTCCTGGCCCTCGCGTGGGCCACA	787
Qy	875	ttgg---gcgcatacaaatgtgaagtlaccggtlccccaagaagagacagagccaaagca	931
Db	788	TCCGACTGCTGCTCTAGGCCCCCGGACACACCCACCTCTTCTTACTCTTGAGACACC	847
Qy	932	agcgaagcccaactctgtatgttggagaataactctacacgctccatgctctccctaa	991
Db	848	AGCTGCTAATGACCTCTATGTGGGGAAGATAAACTGGCTTCTATGCTCTTAAAGAC	907
Qy	992	tgatacagaggggtgtgtcgtgtgcgtgcgcccggcagcaacttcccttgcgtgaagagc	1051
Db	908	TGTGTCCACACAGGATGGTGGCCCTGGTGGCTGTGTGATGACCTGGGCCCGGACGATGGCC	967
Qy	1052	ccgaagactgtagtgcgtacacatcgcggggaagaaggagtgatgtatcaagcccaagag	1111
Db	968	CCACCACAGATGAGGTACACTCCAGTCTCAAGGAAGCGAAGGGCTTACCACACTG	1022
Qy	1112	acgtcaagttgtatcccggaactcaaaagcaagaacaagctcaactacttgaggaattact	1171
Db	1028	CTGTTAATATACCCCTCAGGC-----AGTGTGGCCCTCCCAACCAAGT	1068
Db	1172	ggcttcgtatagacacatgaacaccocactgtctgcgtctcaagaatgctcgggagaat	1231
Db	1070	GGGTGCTCATTTGGACACACGAGACTACCCCACTCTGACACACACATGCTGAAGGTCC	1128
Qy	1232	ttcccaacaatctcccacaacatcgcgaaatgtgatlctcgtctgatltaagagaaaga	1291
Db	1130	ATCCCAACCTGGGAGTGGAACTGACAGACAAAGACTCCAGAAATACCAGGCCAC	1189
Qy	1292	gctttgaggaagtatcaaacctggtgtgaccagactcagaaaaagcaactaccacggt	1351
Db	1190	CTTT-----CTTTTGGACTATTGAGCTGTGAGCCGAGAAAGAACTTTGGACTCCGAGC	1243
Qy	1352	ctcggagatgtgagaaagccgcgcacatgcocctgtcccgccgcggagcccccgtgact	1411
Db	1244	TGC-----ATCCAAACAACAAAAAATCCAGACTTTACTTTGGGGCTGGACCCC-----	1291
Qy	1412	ccatgtctaaagacatggtactacatcaatccctgagcaacttccctgctatgtgcttgggtg	1471
Db	1292	-----AAGACCTCTCTGGGAGCTAGCCTCACTCTCTCTCTCCGAGGAGGTGATTC	1342
Qy	1472	ccctcatcatcaactatccctctgagtgtatcagtcagtcagtcagtcagtcagtcagtcagtc	1531
Db	1343	TCCTTTGATAG-----	1351
Qy	1532	tccagaaggaactgtgagaaatccacgtctctgagcagcagcagtcagtcagtcagtcagtcagtc	1591

Db	1356	-----GAGGTGGTGGAGAAAGCAGACGAGACCC	1387
Qy	1592	accaccctcggagacacgctcctagaacgycgagctcctcgtgacacgctcgtgcctgaatcaag	1651
Db	1388	TGGCACCTCGACGACTTGGCTCA-----	1409
Qy	1652	agagctcgtgacacagacagacccacacgctcccccagggcctccaaacacctcgtcgtc	1711
Db	1410	-----CATCTCCAGATGCGCCAGTCCCTGACACTCGGGGGCCAGCCGGAGGCCAATA	1463
Qy	1712	ccggagactctcgtcctccaagcctcgtcagcagccccctccctcgtgaacaacacgactggagatg	1771
Db	1464	GAGCGTTCAAGTCTCCCTCAAAAGCAAGCCACGCACATC-----GACGACCTCTAAGCTG	1516
Qy	1772	aggaaacacgactgtgtatagatgttggaanaatlctcctcctcgtcccaagatcgtcctggcc	1831
Db	1517	AGCAACTCCACG-----TAGTGGGGAAGATTTCTTCAATCCCAAGGACGTGTGGGCC	1570
Qy	1832	atggagctgagggacacatgtgtgacccgggacatgtttgacaacccgacgctgagccgga	1891
Db	1571	GGGGGGCAGGGGGACCTTTCGTTTTCCGGGGACAGTTTGAGGACCGGANGGCTGTCTCA	1633
Qy	1892	agagagactcccccagactgtgttttagctctgcagacacgctgaagctcagctgtgtcgaagat	1951
Db	1631	AGCGCTCTCCCGGAGTGGCTTGCGTGGGGGGAAGTTCAACTGTCTGAGGAGT	1690
Qy	1953	cggatgagaccacgaacgtgatccgctactctctgcacggaagaaagacccggaatccag	2011
Db	1691	CTGACAGGACACCCACAGGTGCTCGCTACTCTTGCAACCGAGGGACCCAGTTTCACT	1750
Qy	2012	acattgcatcagagctgtgtgcacacacccctgcagaaagatagtgtgagcagaagagactgtg	2077
Db	1751	ACATTGCGCTTGAGACTCTGCCGGGCCCTCTTGCAAGAGTACGTAGAAAACCCGACCTGG	1810
Qy	2072	cgcactcgcgcgtggaagccatcaactctgtcgcagaaacacacccctggacctgcaccac	2131
Db	1811	ATCGGGGGGTCCTGAGCGCCAGAGTCTGCTGACGACGTGATGTCTGGCTCGGCCAAC	1870
Qy	2132	tccactcctcaacatcgtctcacagaagacctaagacacacaacatcctcatatccatc	2191
Db	1871	TGCACCTCTTACACTAGTAGTACCGGAGCCTGTGAAGCAGGAAMATATTCATCACCGGGC	1933
Qy	2192	cbaatgcacacgacgaagatacaagacatgactcctcgacttgcgcctcgtcagaagactg	2251
Db	1931	CTGACAGCCAGAGGCGCTGGGCGAGATGGTGTCTTCAACTTGGCCCTCTCAAGAAGCTGC	1990
Qy	2252	cagtgaggacacaaatgttaagccgcgcgactcgtgtgggtgtcctgtgcacaagaagcttgatgc	2311
Db	1991	CTGCTGGCGCTGTAGCTTCAAGCTTCACTCCACTCCGACATCCCGGACAGGAAGCTGGATGG	2050
Qy	2312	ctccagagaatcctgacgagaagacatgtaagagaagacctaactcaacagcgtgagacatctt	2371
Db	2051	CGCCGAGACTTCTGAGACTCTCTGCACACCAAGTACTTACACAGCGCTGAGCACTTCT	2110
Qy	2372	ctgacagctcgtcttcttactaagctgtgtctcgtgagggcagacaccccttltgtcaagctcc	2433
Db	2111	CTGCAAGCTGTGGTCTTACTACCTGTTTCTGTGTGGCACCAACCCCTTTTGAAGACATC	2170
Qy	2433	tbcacgcgacgaacacaactcctccctgggtgcctcgcagcccttgactcgtgcaccacaga	2491
Db	2171	TTTATTCGGCAGGCAAACTCTCTACAGGGGCTCCCTGTGTGCTACACTGTGAGGAAGAGG	2233
Qy	2492	adcaacgaacgcatcgtacacgaatltgatagaagaatgtatcgtacgtatgactccaga	2551
Db	2231	TTCACAGAACAGGTGTGTGGCCCGGAGACCTGTGTGGAGCCATGTGTGAGCCCACTGCGCAGC	2290
Qy	2552	aacgcacctcagcgaacgactgttcaaaacccgtcttctcgtgaagctcagaagacgacg	2611
Db	2291	CACGCCCTCTGCCCCCAGGTGCTGGCCCAACCCCTTTCTTTTGTGAGCAAGACCAAGCAAC	2350
Qy	2612	tccagttcttcagagactgtagcagacgaatagaanaagaatccctcgtatgtgcgcctatgc	2671
Db	2351	TCCACTTTTTCACAGAGCTCATGTACTGTGGTGAAGAGTGTCCGACAGAGACCCCTGG	2410

OY	1918	agagagccaaagagcagatgacgcccacactctgatatgtgttggaataactcttacacgctctca	977
Db	854	TTCCCTTGACACCCACGAGCTTTTGTGTAACACTGTAATGTGGGGAAAGAAAGAACTCGGGTTTGA	913
OY	978	tgactctccctcacaatgtgtacacgaagggtgtctgtctgtccgcccgcgcgcacacacttcc	1037
Db	914	TGTCTCTAAAGCACTGGTGTCCAGCGCTGGGGGTGGCCCTGCGGCCGTGGAGCTGACCTTGCC	973
OY	1038	tttctgtgaagggcccacagactctgacgtgtcatccatctcgggaacaaagggtgtgtgtat	1097
Db	974	ACCCATGGACAGCTCCCAACACAGATGAGGTGACACTTCAAGCTCTGGGGGAGCCGAGAGGG	1033
OY	1098	cacgcgcagcagcgagcgtacgaattgtatcccgagactcaaaagaagaacaagctcctaeta	1157
Db	1034	CTCACCCAGCACTCTCTTCAGATACCCCTTAGGACAGCTGGCCCTCCAGGC-----	1086
OY	1158	cttgaggaattactcgtctctctgtatagacaccactgaaccccacactgtctgtctaccaa	1217
Db	1087	-----AGTGGCTACTCATTTGGATACACGAGAACTCCCCCGTTTACACACACAC	1135
OY	1218	gatgctgtgagagatctcccaacaactctacccaacactcgggaanaatlgatctctgtca	1277
Db	1136	CATGCTGGGGGTTCATTCCCATC-----CCAGGGAAAGTGTGCTGAGAC	1180
OY	1278	ctcagagaaagagcgtcttgaggaagtatcaacagctgtgcagactctcaagaaaagc	1337
Db	1181	ACGAGCCTCAGAGGAACTCCA-----TGCTCCACCTGTCTTTTGTAGCTCTTAAACTG	1235
OY	1338	acctaccacgcgtgtcctgggagtgtgagagagaagccgcaccactgcctgtcccgcgcga	1397
Db	1236	AGGAGGGAGATCCAGAACTACATTCCTGAAGAAAGCTTCAGACTCTTATCCAGGGCTG	1295
OY	1388	ggcccccgltgactcactgtcttaagacatgtgtacatcatctctgagaccttctgtct	1457
Db	1296	GGATGCC-----AAAGCTCGTAGCAGCTACTTCTACCTGCATCTCTCT	1339
OY	1458	gattgtcgtgggttgcccttcaatcaactatccctcgtagcatgataagaaagacagact	1517
Db	1340	GGGAGCTGGGTCTCTTAACT-----GATGAGGACAGACAGACAGAG	1381
OY	1518	ccagcacccagcagltccagaagaagactgtgagaagatccacgtctctgtcacgcagcagca	1577
Db	1382	CC-----CTTCACGA	1391
OY	1578	gcagctgccttccacccacacttgagacaacgctcaggaagcgagctctctgtgacagctc	1637
Db	1392	CTGTGTGCCCTCTCTGACTCTACAGAGATGCTCAAGGACACACTCTCAAGGACATCTTGA	1451
OY	1638	tggtcccgtaactcagaagcgtcgcggacacagcccacagcaagctcccccagagctccaa	1697
Db	1432	CAGGACAGAGAGAAATTTCAAGCCCTTCAGACCAGCCACCACT-----	1499
OY	1698	ccactcgtctgtccgcgcagcctctgcctcaagsgltgcagcagaccctccctctgnaaca	1757
Db	1500	-----CA	1501
OY	1758	agacgaatgagatgtaggaanaacgcgtgtgtgatgttttggaanaatttccttctgccaa	1817
Db	1502	TGACCTCTGAACACAGGCAAGCCACACCGGTGT-----GGGAAGATTTCCTTCAMCCCAA	1555
OY	1818	ggaatgtctgggcacatgagctctgagggacaaatltgtatccgggcagatglttgaaacgc	1877
Db	1556	GGATGTGCTGGCCGTGGGCAAGAGGAGATTTGTTTCCAGGACAACTTTGAAGGGCG	1615
OY	1878	cgaagctgcgcgtlgaagagagatctcccccagtgltttttagcttgcagacagctggaagtcca	1937
Db	1616	GGCAGTACCTGTGAAGAGGCTCTCTCGAAGAGTCTTTGGCTTGTTCCGCGAGAGAGTGCA	1675
OY	1938	gctgtgtgcagaaatcggatgtagcaaccgcgaactgtatccgtactctctgcacggagaaga	1997
Db	1676	GCTGTGTGAAGAGTCTAAGCAGGACCCCAACAGTGTCTGCTACTCTGACATGACGACATGG	1735

QY	1998	cgaggaattccgaagacatttcacatccgaagcgtgtgtgaagccacccttcgaagaagtattg	2057
Db	1736	ttcccgatttccactactacatttgccctttggagctctccagccctctcttgcaagaaactagctga	1795
QY	2058	gcggaagaagacttttgcgatactgcgtgcgtgagcccatcaacttgcgttcgagcagaccactc	2117
Db	1796	gagccacagactttagatgcgtggggccgtggagccaccacacagatgctgacagcatgatgtct	1855
QY	2118	gggccttgcgcacactccaactccctcaacatcglttccagagaactaaagccacacaaat	2177
Db	1856	cggccttgcccccattctacatttcctttacacatgagttacacgtgaccttmaaccagccacat	1915
QY	2178	cctcatatcaatgaccgaatgcgaacgaagaatcagaagccatgactccgacttgcgact	2237
Db	1916	ttctcatgtgctggcctctacacagccagggccacagccagaaagtgtcatctctgacttggcctt	1975
QY	2238	ctgcgaagaagcttgcgcagtggtgcagacacagtttccagccgcgcatctggtgtgcctgcac	2297
Db	1976	ctgcgaagaagcttaccctgtttggcgtttgtacttcatcttccattctcgtgatttccgggac	2035
QY	2298	agaagcttgcatacgtccccagaagatgctgagcgaagactgtaaggagaacctactaac	2357
Db	2036	agaaagctgcgattgacacacagacagctctgc---cagctccccccagacagccacacacagc	2092
QY	2358	gltggaacatctttctgcagcgtgcgcgtcttctactcagcgtgtgtcctctgagggagacccc	2417
Db	2093	tttggaattttcttccacagagctccgattttttatgctgttccgggtgcagccacc	2152
QY	2418	ttttgcgaagctcccttgcagcgcgcgcacacacccctcgtgtgcctgcgacttgc	2477
Db	2153	cttttgagagagctcttctacacgacagccacacatctctcagggagatccctgtctgctca	2212
QY	2478	cttgcacccagaagaagcagaaagcgtcatctgcagagaaattgataagagaatgattgc	2537
Db	2213	gctgcgaggaamaagaccatgacaaaggttgctgacattgagctttagtagagaccattctag	2272
QY	2538	gatgaatcccttaagaagaagcccttcgaagaaacgcgtgcacaaagccgtcttcttgag	2597
Db	2273	ctttgctgcccacagatgcccccttcctggcagagatggctactgcctcacccctcttttgag	2332
QY	2598	cctagaagaagcagcttccagttcttccagagcgtgcagcgaacagatagaagaagaaatccct	2657
Db	2333	cagagcttaaggaagctccaggttcttccagatgacatgacgtggttgagaaagaaacacga	2392
QY	2658	ggaatgcccgaatcgltgaagcagttagaagagaagcggagagaagccgtgtgtgaagaatgactg	2717
Db	2393	tcagagggcctcttcgtctgcgtgcggcgtgcgagcagaaacctcaagagtgctgtccgggaagactg	2452
QY	2718	gcgaggagaacataactgaacccctccagaagaagcctgcgttaattcagactataaagg	2777
Db	2453	gcacaaagacatcttcaccccctctgcagggagatgcgaaaggttccgtctatcacaaagg	2512
QY	2778	tgattctctcgaagatcttccctccgagcagcagaaataagaagacacactacccggaact	2837
Db	2513	gacatcactttcgaaacctgcttccggcccatagaaacaaagaaacacactacagggagact	2572
QY	2838	gcttcgacaagcttgcggagaaacgtctgggaacccctcccgacgaactcgtgtgtactaac	2897
Db	2573	cccactctagctgcgcagacactagagccaaattctcgtcgtgcttcatctacacttacttccac	2632
QY	2898	gctctgccttcccccaactctctgcgaacacactacccgggcctcagtgagctgtgtgaagcaaga	2957
Db	2633	acagaggaatttccacagagctgtgtgtgcacacactacacgtccatgagacacttgcgctctga	2692
QY	2958	gagactctccagccctactac	2979
Db	2693	gagcctcttctcctgcccctactac	2714

ACCESSION AC016784
VERSION AC016784.2 GI:9123797
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 70508)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome, clone RP11-9J12
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 70508)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galligan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lenoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6538797.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3011
Center clone name: 9_J_12

* NOTE: This record contains 80 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 785 884: contig of 784 bp in length
* 885 gap of 100 bp
* 1668 1767: contig of 783 bp in length
* 1768 2539: contig of 772 bp in length
* 2540 2639: gap of 100 bp
* 2640 3432: contig of 793 bp in length
* 3433 3532: gap of 100 bp
* 3533 4336: contig of 804 bp in length
* 4337 4436: gap of 100 bp
* 4437 5222: contig of 786 bp in length
* 5223 5322: gap of 100 bp
* 5323 6105: contig of 783 bp in length
* 6106 6205: gap of 100 bp
* 6206 6979: contig of 774 bp in length
* 6980 7079: gap of 100 bp
* 7080 7856: contig of 777 bp in length
* 7857 7956: gap of 100 bp
* 7957 8730: contig of 774 bp in length
* 8731 8830: gap of 100 bp

* 8831 9619: contig of 789 bp in length
* 9620 9719: gap of 100 bp
* 9720 10507: contig of 768 bp in length
* 10508 10607: gap of 100 bp
* 10608 11371: contig of 764 bp in length
* 11372 11471: gap of 100 bp
* 11472 12270: contig of 799 bp in length
* 12271 12370: gap of 100 bp
* 12371 13157: contig of 787 bp in length
* 13158 13257: gap of 100 bp
* 13258 14040: contig of 783 bp in length
* 14041 14140: gap of 100 bp
* 14141 14921: contig of 781 bp in length
* 14922 15021: gap of 100 bp
* 15022 15780: contig of 759 bp in length
* 15781 15880: gap of 100 bp
* 15881 16639: contig of 759 bp in length
* 16640 16739: gap of 100 bp
* 16740 17490: contig of 751 bp in length
* 17491 17590: gap of 100 bp
* 17591 18373: contig of 783 bp in length
* 18374 18473: gap of 100 bp
* 18474 19281: contig of 808 bp in length
* 19282 19381: gap of 100 bp
* 19382 20157: contig of 776 bp in length
* 20158 20257: gap of 100 bp
* 20258 21042: contig of 785 bp in length
* 21043 21142: gap of 100 bp
* 21143 21954: contig of 812 bp in length
* 21955 22054: gap of 100 bp
* 22055 22839: contig of 785 bp in length
* 22840 22939: gap of 100 bp
* 22940 23724: contig of 785 bp in length
* 23725 23824: gap of 100 bp
* 23825 24606: contig of 782 bp in length
* 24607 24706: gap of 100 bp
* 24707 25472: contig of 766 bp in length
* 25473 25572: gap of 100 bp
* 25573 26362: contig of 790 bp in length
* 26363 26462: gap of 100 bp
* 26463 27247: contig of 785 bp in length
* 27248 27347: gap of 100 bp
* 27348 28113: contig of 766 bp in length
* 28114 28213: gap of 100 bp
* 28214 29011: contig of 798 bp in length
* 29012 29111: gap of 100 bp
* 29112 29898: contig of 787 bp in length
* 29899 29998: gap of 100 bp
* 29999 30779: contig of 781 bp in length
* 30780 30879: gap of 100 bp
* 30880 31679: contig of 800 bp in length
* 31680 31779: gap of 100 bp
* 31780 32633: contig of 854 bp in length
* 32634 32733: gap of 100 bp
* 32734 33490: contig of 757 bp in length
* 33491 33590: gap of 100 bp
* 33591 34362: contig of 772 bp in length
* 34363 34462: gap of 100 bp
* 34463 35229: contig of 767 bp in length
* 35230 35329: gap of 100 bp
* 35330 36093: contig of 764 bp in length
* 36094 36193: gap of 100 bp
* 36194 36953: contig of 760 bp in length
* 36954 37053: gap of 100 bp
* 37054 37857: contig of 804 bp in length
* 37858 37957: gap of 100 bp
* 37958 38752: contig of 795 bp in length
* 38753 38852: gap of 100 bp
* 38853 39662: contig of 810 bp in length
* 39663 39762: gap of 100 bp
* 39763 40549: contig of 787 bp in length
* 40550 40649: gap of 100 bp
* 40650 41424: contig of 775 bp in length